

Data-Driven Distributionally Robust Electric Vehicle Balancing for Mobility-on-Demand Systems under Demand and Supply Uncertainties

Sihong He Lynn Pepin Guang Wang Desheng Zhang Fei Miao

Abstract—As electric vehicle (EV) technologies become mature, EV has been rapidly adopted in modern transportation systems, and is expected to provide future autonomous mobility-on-demand (AMoD) service with economic and societal benefits. However, EVs require frequent recharges due to their limited and unpredictable cruising ranges, and they have to be managed efficiently given the dynamic charging process. It is urgent and challenging to investigate a computationally efficient algorithm that provide EV AMoD system performance guarantees under model uncertainties, instead of using heuristic demand or charging models. To accomplish this goal, this work designs a data-driven distributionally robust optimization approach for vehicle supply-demand ratio and charging station utilization balancing, while minimizing the worst-case expected cost considering both passenger mobility demand uncertainties and EV supply uncertainties. We then derive an equivalent computationally tractable form for solving the distributionally robust problem in a computationally efficient way under ellipsoid uncertainty sets constructed from data. Based on E-taxi system data of Shenzhen city, we show that the average total balancing cost is reduced by 14.49%, the average unfairness of supply-demand ratio and utilization is reduced by 15.78% and 34.51% respectively with the distributionally robust vehicle balancing method, compared with solutions which do not consider model uncertainties.

I. INTRODUCTION

There are over 5 million EVs by December 2018 globally, and this figure is predicted to increase to 100-125 million by 2030 [3]. Compared to conventional gas vehicles, EV fleets have prolonged charging time and concentrated mobility patterns due to current charging technologies and limited charging infrastructures, especially for commercial EV fleets, e.g., e-taxi, future autonomous mobility-on-demand (AMoD) systems, given their long daily travel distances [14].

Researchers have been focusing on models and algorithms to study EVs [16], [14]. There have also been

This work is partially supported by NSF SAS-1849246 and NSF CPS-1932250. Sihong He, Lynn Pepin, and Fei Miao are with the Department of Computer Science and Engineering, University of Connecticut, Storrs Mansfield, CT, USA 06268. Email: {sihong.he, lynn.pepin, fei.miao}@uconn.edu. This work is also partially supported by NSF SAS-1849238 and CPS-1932223. Guang Wang and Desheng Zhang are with the Department of Computer Science, Rutgers University, Piscataway, NJ, USA 08901. Email: {gw255, desheng.zhang}@cs.rutgers.edu.

focusing on how to choose the optimal locations for charging stations and how to assign charging points to EVs in each station to minimize the charging time of EVs considering various constraints, e.g., demand, costs, and charging compatibility [6], [15]. However, high costs of charging infrastructures and land resources make it impractical to deploy abundant charging stations and points at the early promotion stage [15]. Even when there is enough charging infrastructure for all EVs in theory, the uncontrolled and decentralized charging and mobility behaviors of some EV fleets, e.g., e-taxi, cause long waiting times when the demand for charging points greatly exceeds the availability [12].

The above mentioned EV management issues have posed key optimization and scheduling algorithm challenges for world-wide EV adoption of AMoD. The interaction between AMoD systems and power networks through EVs based on the vehicles' charging requirements, battery depreciation, and power transmission constraints have been investigated, and the economic and societal value of EV AMoD has been analyzed [11]. To improve the performance of general AMoD systems, mobility demand based vehicle balancing methods have been proposed with various system design objectives, such as reducing the number of vehicles needed to serve all passengers [19], [4], reducing customers' waiting time [13], or taxis' total idle distance [8]. However, the limited knowledge we have about charging patterns [14] affect the performance of vehicle balancing strategies, and make real-time decisions under demand model uncertainties still a challenging and unsolved task.

The contributions of this work are as follows:

- We are the first to consider both future demand uncertainties and EV supply uncertainties predicted based on charging activity data in designing a system-level vehicle balancing algorithm. While model predictive control algorithms [19], [4] have been designed considering AMoD system demand uncertainties in the literature, the supply side uncertainties for EV AMoD is not well studied yet.
- We design a distributionally robust optimization approach to balance EVs across a city for minimum total idle distance and balanced charging

station utilization with respect to the worst-case expected cost. The approach considers probability distribution uncertainties of the passenger mobility demand and the EV supply caused by the challenge of charging process prediction [15], [12].

- We derive an equivalent form of convex optimization problem for the proposed distributionally robust optimization problem to provide system-level performance guarantee in a computationally tractable way under model uncertainties. Based on real data of Shenzhen city, we show that the average total balancing cost is reduced by 14.49%, the average unfairness of supply-demand ratio and utilization is reduced by 15.78% and 34.51%, respectively, with the proposed method, compared with solutions which do not consider model uncertainties.

The rest of the paper is organized as follows. The distributionally robust EV balancing problem is presented in Section II. An equivalent computationally tractable form is derived in Section III. We show performance improvement in experiments based on real data in Section IV. Concluding remarks are provided in Section V.

II. PROBLEM FORMULATION

In this section, we formulate a distributionally robust optimization problem to balance EVs across a city with minimum total idle distance and balanced charging station utilization. Both passenger mobility demand and EV supply uncertainties are considered. The region every empty EV will go is updated in a receding horizon control process. At each time step, the EV status is updated to the dispatch center first, then the dispatch center calculates a vehicle balancing decision by solving the proposed distributionally robust optimization problem, and sent solutions to EVs. The goal is to dispatch vacant EVs to different regions to pick up current and predicted passengers if the EVs have enough energy, or to charging stations if the EVs are short of energy, while minimize the cost of dispatching for the following τ time steps. Local dispatchers that match individual EV with one or several passengers (for carpool) is out the scope of this work.

A. EV States and Corresponding Actions

We assume there are three possible states for one EV: vacant, occupied, and low-battery. Vacant means there are no passengers in this EV, and it has enough energy to finish the next trip. The controller dispatches vacant EVs according to current and predicted passengers demand. When a vacant EV picks up one or more passengers, it turns to occupied, and the controller has no actions

for it until it becomes vacant again. An occupied EV will be finishing current order in a time period and will become a vacant EV once it drops off its passengers. One occupied EV can only become a vacant EV when it finishes the current order. When a vacant EV can not finish the next trip with the remaining battery, this EV becomes a low-battery EV and will go to regions assigned by the controller where it can find a charging station. Before a low-battery EV gets fully charged, it stays in the low-battery status until it leaves the charging station and becomes vacant. A low-battery can only transfer to a vacant EV or stay in current state.

B. Problem Description

We assume that one day is divided into K time intervals, and we use $k = 1, 2, \dots, K$ to denote time index. We assume the entire city is divided into N regions and we use n to denote region index, where $n = 1, 2, \dots, N$. At time k , the system-level controller makes vacant and low-battery EVs to go to other regions or stay in the same region for picking up passengers or charging, respectively. After one EV arrives at its dispatched region, a local-level controller assign the EV to pick up passengers or to charge according to the EV's battery status.

During time k , there are r_i^k predicted total amount of passengers demand and c_i^k predicted total number of EVs finish charging (new supply of EVs) in region i , where $i = 1, 2, \dots, N, k = 1, 2, \dots, K$. Let demand vector $r^k = [r_1^k, r_2^k, \dots, r_N^k]^T$ and supply vector $c^k = [c_1^k, c_2^k, \dots, c_N^k]^T \in \mathbb{R}^N$ be random vectors instead of deterministic vectors. And assuming they are independent. To model the spatial and temporal relations of deman(supply) during every τ consecutive time interval, we define concatenation of demand as $r = [r^1, r^2, \dots, r^\tau]$, and concatenation of supply as $c = [c^1, c^2, \dots, c^\tau]$. We use F_r^* and F_c^* to denote the unknown true probability distributions of $r, c \in \mathbb{R}^{N\tau}$ respectively, i.e. $r \sim F_r^*$ and $c \sim F_c^*$.

We use non-negative matrices X^k and Y^k as the decision matrices at time k where $X^k, Y^k \in \mathbb{R}_+^{N \times N}$ and $x_{ij}^k(y_{ij}^k)$ is the total amount of vacant(low-battery) EVs will be dispatched from region i to region j at the beginning of time k . Minimizing the expected allocating cost given true probability distributions of demand vector and supply vector is defined as the following stochastic programming problem:

$$\begin{aligned} \min_{X^{1:\tau}, Y^{1:\tau}} \quad & \mathbb{E}_{r \sim F_r^*, c \sim F_c^*} [J(X^{1:\tau}, Y^{1:\tau}, r, c)] \\ \text{s.t.} \quad & X^{1:\tau}, Y^{1:\tau} \in \mathcal{D}, \end{aligned} \quad (1)$$

where $J(X^{1:\tau}, Y^{1:\tau}, r, c)$ is a cost function of allocating EVs according to decisions $X^{1:\tau} = \{X^1, X^2, \dots, X^\tau\}$ and $Y^{1:\tau} = \{Y^1, Y^2, \dots, Y^\tau\}$ under demand r , supply c and convex constraints domain of decision variables \mathcal{D} .

However, in real world scenarios, we usually have limited knowledge about the true probability distributions of r and c . Though we have historical or streaming data, we can only estimate a set of probability distributions, such that $F_r^* \in \mathcal{F}_c$, $F_c^* \in \mathcal{F}_c$ considering the randomness of the parameters and prediction errors [12], [15], [11], instead of knowing the exactly form of F_r^* and F_c^* . Meanwhile, we notice that problem (1) is computationally expensive to solve. So in this work, we consider to minimize the worst case expected cost function which is a minmax form of problem (1) as below:

$$\begin{aligned} \min_{X^{1:\tau}, Y^{1:\tau}} \max_{F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c} \mathbb{E} [J(X^{1:\tau}, Y^{1:\tau}, r, c)] \\ \text{s.t. } X^{1:\tau}, Y^{1:\tau} \in \mathcal{D}. \end{aligned} \quad (2)$$

Problem (2) is a form of distributionally robust optimization problem [2] and assume $F_r^* \in \mathcal{F}_c$, $F_c^* \in \mathcal{F}_c$. In the following sections, we will define the complete forms of object function and constraints as well as the probability uncertainty sets $\mathcal{F}_r, \mathcal{F}_c$.

C. Cost of EV Balancing J_D

We aim to balance the EV supply according to passenger mobility demand, by sending vacant state EVs to serve passengers according to X^k and low-battery EVs to charging stations according to Y^k . Given a specific region partition method, let $W \in \mathbb{R}^{N \times N}$ be the cost matrix where w_{ij} is the cost sending a vacant EV from region i to region j . The cost can be metrics such as the approximated distance, the minimal routing distance or travel time between two regions. Here, we use approximated distance from two regions as the dispatching cost. Let $W^* \in \mathbb{R}^{N \times N}$ be the cost matrix of sending one low-battery EV which is partially the same as W . When there is at least one charging station in region j , $w_{ij}^* = w_{ij}$ for all i . If there are no charging stations in region j , $w_{ij}^* = \infty$ for all i since low-battery EV should not go to regions without charging stations. Then the total re-balancing cost function J_D for τ intervals is

$$J_D(X^{1:\tau}, Y^{1:\tau}) = \sum_{k=1}^{\tau} \sum_{i=1}^N \sum_{j=1}^N (x_{ij}^k w_{ij} + \beta y_{ij}^k w_{ij}^*), \quad (3)$$

where β is a weight coefficient. Since the distance EVs can move during a given time interval is limited, we

have the following constraints for variables X^k and Y^k .

$$\begin{aligned} x_{ij}^k \geq 0 \text{ and } x_{ij}^k = 0 \text{ when } w_{ij} \geq m_1; \\ y_{ij}^k \geq 0 \text{ and } y_{ij}^k = 0 \text{ when } w_{ij}^* \geq m_2, \end{aligned} \quad (4)$$

where $m_1 > 0, m_2 > 0$ is the upper bound balancing distance for an vacant and low-battery EV, respectively. These constraints consider real-world scenarios that we cannot dispatch EV to some far away regions when the moving distance exceeds the capability, either due to speed limit or insufficient battery.

D. Utilization of Charging Stations J_E

Given fixed number and locations of charging stations in the city, to avoid the long waiting time at some charging stations, one method is to balance the charging station utilization across the whole city. It will improve charging efficiency of EVs as well as decrease driver's potential cost due to EV's unique charging problems. The supply variable c_i^k is the service rate or the average number of new available charging spots (due to EVs finished charging) in region i during time k . The net total number of low-battery EVs $Y_i^k = \sum_{j=1}^N y_{ji}^k - \sum_{j=1}^N y_{ij}^k$ in region i after balancing according to decision variable Y^k is the average arrival rate or the average number of arriving EVs. The overall utilization of charging stations in region i during time k is approximated as $\frac{Y_i^k}{c_i^k}$. The total difference between the inverse local and inverse global utilization for τ time intervals is

$$\sum_{k=1}^{\tau} \sum_i^N \left| \frac{c_i^k}{Y_i^k} - \frac{\sum_{j=1}^N c_j^k}{\sum_{j=1}^N Y_j^k} \right|. \quad (5)$$

However, the charging station supply is a random variable that the prediction error can not be ignored [5], [12], [15], and we define the uncertainty set as $c \sim F_c^*, F_c^* \in \mathcal{F}_c$. Function (5) is not concave over uncertainty parameter c^k , for computationally tractability, we consider minimizing the following utilization quality function J_E

$$J_E(Y^{1:\tau}) = \sum_{k=1}^{\tau} \sum_{i=1}^N \frac{c_i^k}{(Y_i^k)^a}. \quad (6)$$

According to Lemma 1 in [9], when the power parameter $a > 0$ is designed to be small enough, objective function (6) is linear in c^k and convex in $Y^{1:\tau}$ and approximates the objective (5).

E. Constraints Definitions

1) *Service quality metrics*: Demand-supply ratio is one service quality metric for AMoD systems [19], [9]. In this work we also minimize the total difference between local and global demand-supply ratio for τ time

intervals. Let $V_i^k, O_i^k \in \mathbb{R}_+$ be the number of vacant and occupied EVs respectively at region i at the beginning of time k before balancing, and $V^k, O^k \in \mathbb{R}_+^N$. Define $S_i^k > 0$ as the total number of supply EVs available to be dispatched in region j during time k , dispatch decisions as $X^{1:k} = \{X^1, X^2, \dots, X^k\}$, $Y^{1:k} = \{Y^1, Y^2, \dots, Y^k\}$. Then the following equations of V_i^k, O_i^k, S_i^k describe dynamics for τ time steps:

$$\begin{aligned} S_i^k &= \sum_{j=1}^N x_{ji}^k - \sum_{j=1}^N x_{ij}^k + V_i^k = X_i^k + V_i^k > 0, \\ V_i^{k+1} &= \sum_{j=1}^N P_{vji}^k S_j^k + \sum_{j=1}^N Q_{vji}^k O_j^k + c_i^k, \\ O_i^{k+1} &= \sum_{j=1}^N P_{oji}^k S_j^k + \sum_{j=1}^N Q_{oji}^k O_j^k, \end{aligned} \quad (7)$$

where $X_i^k = \sum_{j=1}^N x_{ji}^k - \sum_{j=1}^N x_{ij}^k$ is the net change of available EVs due to decision variable X^k at region i . $P_v^k, P_o^k, Q_v^k, Q_o^k \in \mathbb{R}^{N \times N}$ are region transition matrices: $P_{vji}^k (P_{oji}^k)$ is the probability that a vacant EV moves from region j at the beginning of time k will transverse to region i and being vacant (occupied) at the beginning of time $k+1$, respectively. Similarly, Q_{vji}^k and Q_{oji}^k are the probability that an occupied EV moves from region j at time k will go to region i and being vacant and occupied at the beginning of time $k+1$, respectively. When receding the time horizon, GPS locations (region information) and status of all EVs will always be updated by real-time sensing data and V^1, O^1 are provided by real-time data. We consider the following service quality constraints to make sure the demand-supply ratio of each region is within a similar range for service fairness:

$$l_i^k \leq \frac{r_i^k}{S_i^k} \leq h_i^k, \quad k = 1, \dots, \tau, \quad (8)$$

where $l_i^k (h_i^k)$ is the lower(upper) bound of the supply-demand ratio in region i at time k . The value of l_i^k and h_i^k are decided by historical data. We transfer the inequalities (8) to the following equations form with slack variables D_i^k, U_i^k :

$$\begin{aligned} r_i^k - l_i^k S_i^k - (D_i^k)^2 &= 0, \\ r_i^k - h_i^k S_i^k + (U_i^k)^2 &= 0, \quad k = 1, \dots, \tau. \end{aligned} \quad (9)$$

2) *Constraints on decision variables $Y^{1:\tau}$* : We define $L_i^k \in \mathbb{R}_+$ be the total amount of low-battery EVs in region i before balancing at the beginning of time k .

Then L_i^k should have the following relationship with S_i^k and Y^k

$$L_i^{k+1} = \sum_{j=1}^N y_{ji}^k - \sum_{j=1}^N y_{ij}^k + \sum_{j=1}^N P_{lji}^k S_j^k > 0, \quad (10)$$

where $L^k \in \mathbb{R}_+^N$ and L^1 is given by real-time data, $P_l^k \in \mathbb{R}^{N \times N}$ is the region transition matrix: P_{lji}^k is the probability that a vacant EV moves from region j at the beginning of time k will go to region i and being low-battery status at the beginning of time $k+1$. Thus the region transition matrices estimated from data satisfy that $\sum_{j=1}^N P_{lij}^k + P_{vij}^k + P_{oij}^k = 1$ and $\sum_{j=1}^N Q_{oij}^k + Q_{vij}^k = 1$.

F. Predicted Model and Uncertainty Set Construction

Instead of assuming we know the true probability distributions of r and c from data, in this work we construct uncertainty sets \mathcal{F}_r and \mathcal{F}_c that describe possible probability distributions of r and c by applying the Algorithm 2 proposed in [10] which constructs distributional sets with a general prediction model. Here we use the autoregressive integrated moving average model (ARIMA model) [18] as the prediction model to capture the spatial and temporal correlations for the predictions. The coefficients of ARIMA model can be estimated by maximum likelihood estimation. We use \hat{r} and \hat{c} to denote the predicted value, \tilde{r} and \tilde{c} to denote the sample value of r and c , respectively. And then we get a corresponding estimation residuals as

$$\tilde{\delta}_r = \tilde{r} - \hat{r}, \quad \tilde{\delta}_c = \tilde{c} - \hat{c}. \quad (11)$$

Then we use these estimation residuals to construct distribution sets \mathcal{F}_r and \mathcal{F}_c that contain the true distribution of r and c with probability at least $1 - \alpha_h$ where α_h is a significant values can be freely chosen from 0 to 1. \mathcal{F}_r and \mathcal{F}_c have the following format:

$$\begin{aligned} \mathcal{F}_r(\hat{r}, \hat{\Sigma}_r, \hat{\gamma}_{1r}, \hat{\gamma}_{2r}) &= \{r = \hat{r} + \delta_r : (\mathbb{E}[\delta_r])^T \hat{\Sigma}_r^{-1} \mathbb{E}[\delta_r] \leq \hat{\gamma}_{1r}, \\ &\quad \mathbb{E}(\delta_r \delta_r^T) \leq \hat{\gamma}_{2r} \hat{\Sigma}_r\}; \\ \mathcal{F}_c(\hat{c}, \hat{\Sigma}_c, \hat{\gamma}_{1c}, \hat{\gamma}_{2c}) &= \{c = \hat{c} + \delta_c : (\mathbb{E}[\delta_c])^T \hat{\Sigma}_c^{-1} \mathbb{E}[\delta_c] \leq \hat{\gamma}_{1c}, \\ &\quad \mathbb{E}(\delta_c \delta_c^T) \leq \hat{\gamma}_{2c} \hat{\Sigma}_c\}, \end{aligned} \quad (12)$$

where $\delta_r (\delta_c)$ is the difference between true value $r(c)$ and predicted value $\hat{r}(\hat{c})$. Since we don't have true value of the difference $\delta_r = r - \hat{r}$ ($\delta_c = c - \hat{c}$), we use estimation residuals to capture the information from these difference. $\hat{\Sigma}_r (\hat{\Sigma}_c)$ is the estimated covariance, $\hat{\gamma}_{1r}, \hat{\gamma}_{2r}$ ($\hat{\gamma}_{1c}, \hat{\gamma}_{2c}$) are two estimated threshold values of $\delta_r (\delta_c)$ by Algorithm 2 in [10] based on the concept of bootstrapping. More discussions about uncertainty set construction refer to [2], [10].

G. Distributionally Robust EV Balancing Problem

The final goal is to dispatch EVs under minimal cost. We define a weight parameter θ of the two objectives J_D defined in (3) and J_E defined in (6). With constraints (4), (7), (9), (10), we define the following distributionally robust EVs balancing problem under uncertain probability distributions of random demand and supply:

$$\begin{aligned} & \min_{\substack{X^{1:\tau}, Y^{1:\tau}, S^{1:\tau}, D^{1:\tau}, U^{1:\tau}, \\ U^{1:\tau}, V^{2:\tau}, O^{2:\tau}, L^{2:\tau}}} \max_{\{F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c\}} \mathbb{E}[J_D + \theta J_E] \\ & \text{s.t. (4), (7), (9), (10).} \end{aligned} \quad (13)$$

Since above problem (13) can not be calculated in polynomial time directly, we derive a computationally tractable form of this problem in the following section.

III. COMPUTATIONALLY TRACTABLE FORM

In this section, we derive the theoretical result, Theorem 1 of this work, i.e., a computationally tractable and equivalent convex optimization form for problem (13) via strong duality. Hence, the optimal solution of (13) can be calculated in real time considering both passenger demand and EV supply uncertainties.

Theorem 1. *The distributionally robust optimization problem (13) with two distributional sets (12) is equivalent to the following convex optimization problem*

$$\begin{aligned} & \min_{\substack{X^{1:\tau}, Y^{1:\tau}, D^{1:\tau}, U^{1:\tau}, \\ S^{1:\tau}, V^{2:\tau}, O^{2:\tau}, L^{2:\tau}, \\ Q_r, q_r, v_r, t_r, Q_c, q_c, v_c, t_c}} H_o + v_r + t_r + v_c + t_c \\ & \text{s.t.} \begin{cases} \begin{bmatrix} v_r & \frac{1}{2}(q_r + \lambda_U + \lambda_D)^T \\ \frac{1}{2}(q_r + \lambda_U + \lambda_D) & Q_r \end{bmatrix} \succeq 0, \\ \begin{bmatrix} v_c & \frac{1}{2}(q_c + \lambda_V - Z)^T \\ \frac{1}{2}(q_c + \lambda_V - Z) & Q_c \end{bmatrix} \succeq 0, \\ t_r \geq (\hat{\gamma}_{2r} \hat{\Sigma}_r + \hat{r} \hat{r}^T) \cdot Q_r + \hat{r}^T q_r \\ \quad + \sqrt{\hat{\gamma}_{1r}} \|\hat{\Sigma}_r^{1/2} (q_r + 2Q_r \hat{r})\|_2, \\ t_c \geq (\hat{\gamma}_{2c} \hat{\Sigma}_c + \hat{c} \hat{c}^T) \cdot Q_c + \hat{c}^T q_c \\ \quad + \sqrt{\hat{\gamma}_{1c}} \|\hat{\Sigma}_c^{1/2} (q_c + 2Q_c \hat{c})\|_2, \\ Q_r, Q_c, \lambda, v_r, v_c \succeq 0, \\ Z_i^k \geq \frac{1}{(Y_i^k)^a} \\ x_{ij}^k \geq 0 \text{ and } x_{ij}^k = 0 \text{ when } w_{ij} \geq m; \\ y_{ij}^k \geq 0 \text{ and } y_{ij}^k = 0 \text{ when } w_{ij}^* \geq m, \end{cases} \end{aligned} \quad (14)$$

where $H_o = J_D - (\lambda_S^T f_S + \lambda_O^T f_O + \lambda_L^T f_L + \lambda_S^T S^{1:\tau} + \lambda_L^T L^{2:\tau}) - \lambda_D^T \text{diag}(lS^T - DD^T) - \lambda_U^T \text{diag}(hS^T + UU^T) - \sum_{k=1}^{\tau-1} (-V_i^{k+1} + \sum_{j=1}^N P_{vji}^k S_j^k + \sum_{j=1}^N Q_{vji}^k O_j^k) \lambda_{V_i^k}$, J_D is defined as (3).

TABLE I

THRESHOLDS $\hat{\gamma}_{1c}$ AND $\hat{\gamma}_{2c}$ FOR DIFFERENT SAMPLE NUMBER N_B

N_B	10	20	50	100	500	1000
$\hat{\gamma}_{1c}$	1.504	0.964	0.576	0.399	0.296	0.176
$\hat{\gamma}_{2c}$	3.715	2.832	2.006	1.768	1.374	1.317

Proof. See Appendix VI-A.

IV. EVALUATION WITH REAL-WORLD E-TAXI DATA

In this section, we evaluate the performance of the proposed distributionally robust optimization-based EV balancing algorithm (1) with three-week E-taxi data from the Chinese city Shenzhen (one of the largest cities in China that operates over 10,000 E-taxis).

In total, there are four different datasets used in this paper, including E-taxi GPS data (vehicle ID, locations, time and speed, etc), transaction data (vehicle ID, pick-up and drop-off time, pick-up and drop-off location, travel distance, etc), charging station data (locations, name, the number of charging points, etc), and the urban partition data (geographic boundaries of 491 small separate regions composing Shenzhen).

We first utilize a widely adopted spatiotemporal constraint-based method [6], [15], [12] to extract charging events of E-taxis and re-split Shenzhen into 54 large areas based on these data. After obtaining the charging events, we utilize the first two weeks as the training data to determine the uncertainty sets and parameters. The remaining one week is used as the testing data to compare the vehicle balancing cost.

A. Uncertainty Set

Here we select 5 busy regions' data as the input for Algorithm 2 in [10] to construct uncertainty sets. We set time horizon τ as 2 and a significant $\alpha_h = 0.25$ to make sure the true probability distributions of demand r and supply c are separately contained in the constructed uncertainty sets with probability at least 75%. We show how thresholds $\hat{\gamma}_{1c}$, $\hat{\gamma}_{2c}$ change for different sample in table I. The value of $\hat{\gamma}_{1c}$ and $\hat{\gamma}_{2c}$ decrease when N_B increase and the speed of decreasing becomes slow as N_B becoming larger and larger. We also notice that as N_B increase, the corresponding estimated moments have a trend to converge to a certain constant which meets bootstrapping algorithm's intuition.

B. Predicted Model

As mentioned in section II-F, we use ARIMA model [18] to predict r and c where current values are predicted by former concatenation of values. Fig. 1 compares ARIMA model's predicted values and historical values of c and r on one day. We can see that ARIMA

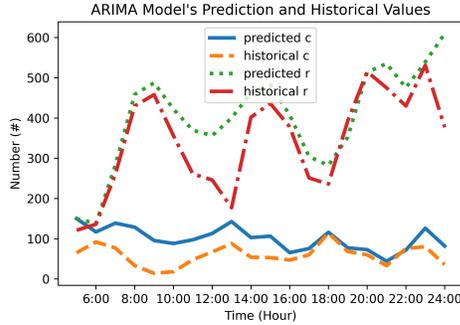


Fig. 1. ARIMA model demonstrates time trends very well.

model demonstrates data's time trends very well and the predicted values of r in peak hours: 8am-10am, 2pm-4pm, 8pm-10pm, are very close to historical values.

C. The Performance of the Proposed Method

We compare our distributionally robust method with the non-robust method [17] by using the same real-time sensing data to evaluate the performance. In the robust method we use the demand and supply uncertainty sets defined as (12), while in the non-robust method, they are deterministic vectors. In Fig 2, we compare the total driving distance of applying the optimal decision of each method from 5am to next day's 12am, 19 hours in total. Here, the total driving distance J_D in (3) is defined as a weighted sum of the charging idle distance and the service idle distance for one EV. Total driving distance is supposed to be small since it implies low total cost. We can see that most of time, the total cost of the robust model is lower than that of the non-robust model. In particular, the average total driving distance is reduced by 14.49% compared with the non-robust method.

In Fig 3 and 4, we compare the unfairness of supply-demand ratio and utilization of the whole city taking decisions of each method. The unfairness metric of utilization (supply-demand ratio) is designed as the total sum of absolute difference between the inverse-local utilization (local supply-demand ratio) and inverse-global utilization (global supply-demand ratio). Lower sum of absolute difference means higher fairness but lower unfairness. The unfairness of utilization in (5) is actually accurate version of J_E in (6) before we approximate J_E . The unfairness of supply-demand ratio in section II-E.1 is the service quality metric for taxi dispatch in [7]: the lower the service quality metric, the better customers' accumulative satisfaction is since they can enjoy uniform quality service. By using robust optimization, the average unfairness of supply-demand ratio and utilization is reduced by 15.78% and 34.51% respectively compared to NonRobust method.

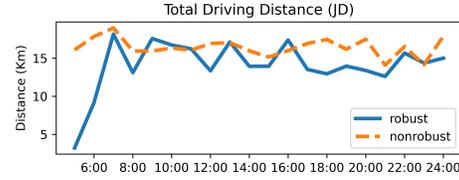


Fig. 2. By using robust optimization, the average total driving distance is reduced by 14.49% compared to NonRobust method.

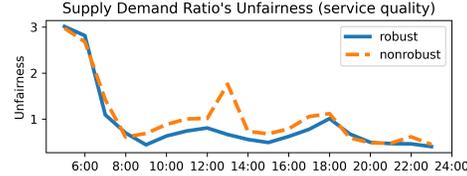


Fig. 3. By using robust optimization, the average unfairness of supply-demand ratio is reduced by 15.78% compared to NonRobust method.

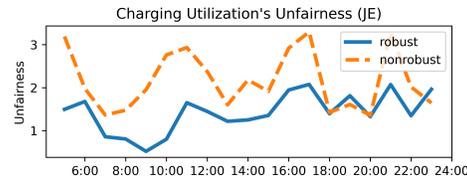


Fig. 4. By using robust optimization, the average unfairness of utilization is reduced by 34.51% compared to NonRobust method.

V. CONCLUSION

Autonomous mobility-on-demand systems can provide more efficient services, and the total idle distance can be reduced with vehicle balancing algorithms in general. However, with the increasing amount of EVs and the limited charging facilities in the city, the uncertainty of charging time (including waiting time) at a charging station affects the EVs supply to provide efficient service for AMoD systems. In this paper, we design a data-driven distributionally robust EV balancing method to minimize the worst-case expected cost under uncertainties about the probability distributions of both demand and supply. Besides reducing EVs' total idle driving distance, we also balance the demand-supply ratios and the charging station utilization of different regions among the city. Then we prove an equivalent computationally tractable form of the distributionally robust problem under the ellipsoid uncertainty sets constructed from data. Evaluations based on three-week real-world E-taxi data from the Chinese city Shenzhen show that the average total balancing cost is reduced by 14.49%, the average unfairness of supply-demand ratio and utilization is reduced by 15.78% and 34.51%, respectively. In the future, we will further evaluate our algorithm based on large-scale data of several years from multiple cities.

REFERENCES

- [1] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, USA, 2004.
- [2] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612, 2010.
- [3] B. N. E. Finance. Electric buses in cities: Driving towards cleaner air and lower CO_2 . http://www.c40.org/c40_research, 2018.
- [4] R. Iglesias, F. Rossi, K. Wang, D. Hallac, J. Leskovec, and M. Pavone. Data-driven model predictive control of autonomous mobility-on-demand systems. In *IEEE International Conference on Robotics and Automation*, volume abs/1709.07032, 2018.
- [5] R. Iglesias, F. Rossi, R. Zhang, and M. Pavone. A bcnp network approach to modeling and controlling autonomous mobility-on-demand systems. *The International Journal of Robotics Research*, 38(2-3):357–374, 2019.
- [6] Y. Li, J. Luo, C.-Y. Chow, K.-L. Chan, Y. Ding, and F. Zhang. Growing the charging station network for electric vehicles with trajectory data analytics. In *Data Engineering (ICDE), 2015 IEEE 31st International Conference on*, pages 1376–1387. IEEE, 2015.
- [7] F. Miao, S. Han, A. M. Hendawi, M. E. Khalefa, J. A. Stankovic, and G. J. Pappas. Data-driven distributionally robust vehicle balancing using dynamic region partitions. In *Proceedings of the 8th International Conference on Cyber-Physical Systems, ICCPS '17*, pages 261–271, New York, NY, USA, 2017. ACM.
- [8] F. Miao, S. Han, S. Lin, J. A. Stankovic, H. Huang, D. Zhang, S. Munir, T. He, and G. J. Pappas. Taxi dispatch with real-time sensing data in metropolitan areas: A receding horizon control approach. *IEEE Transactions on Automation Science and Engineering*, 13:463–478, April 2016.
- [9] F. Miao, S. Han, S. Lin, Q. Wang, J. A. Stankovic, A. Hendawi, D. Zhang, T. He, and G. J. Pappas. Data-driven robust taxi dispatch under demand uncertainties. *IEEE Transactions on Control Systems Technology*, 27(1):175–191, Jan 2019.
- [10] F. Miao, S. He, S. Han, A. M. Hendawi, M. E. Khalefa, J. A. Stankovic, and G. J. Pappas. Data-driven distributionally robust optimization for vehicle balancing of autonomous ride-sharing systems. *CoRR*, abs/2030.61468, 2020.
- [11] F. Rossi, R. D. Iglesias, M. Alizadeh, and M. Pavone. On the interaction between autonomous mobility-on-demand systems and the power network: models and coordination algorithms. *IEEE Transactions on Control of Network Systems*, pages 1–1, 2019.
- [12] Z. Tian, T. Jung, Y. Wang, F. Zhang, L. Tu, C. Xu, C. Tian, and X. Y. Li. Real-time charging station recommendation system for electric-vehicle taxis. *IEEE Transactions on Intelligent Transportation Systems (TITS)*, 17(11):3098–3109, 2016.
- [13] A. Wallar, M. Van Der Zee, J. Alonso-Mora, and D. Rus. Vehicle rebalancing for mobility-on-demand systems with ride-sharing. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4539–4546, Oct 2018.
- [14] G. Wang, X. Chen, F. Zhang, Y. Wang, and D. Zhang. Experience: Understanding long-term evolving patterns of shared electric vehicle networks. In *Proceedings of the 25th annual international conference on Mobile computing and networking (MobiCom)*. ACM, 2019.
- [15] G. Wang, W. Li, J. Zhang, Y. Ge, Z. Fu, F. Zhang, Y. Wang, and D. Zhang. sharedcharging: Data-driven shared charging for large-scale heterogeneous electric vehicle fleets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, 3(3):108, 2019.
- [16] G. Wang, X. Xie, F. Zhang, Y. Liu, and D. Zhang. bcharge: Data-driven real-time charging scheduling for large-scale electric bus fleets. In *2018 IEEE Real-Time Systems Symposium (RTSS)*, pages 45–55, Dec 2018.
- [17] Y. Yuan, D. Zhang, F. Miao, J. Chen, T. He, and S. Lin. p2charging proactive partial charging for electric taxi systems. In *IEEE International Conference on Distributed Computing Systems, ICDCS'19*, 2019.
- [18] G. Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50(Supplement C):159 – 175, 2003.
- [19] R. Zhang, F. Rossi, and M. Pavone. Model predictive control of autonomous mobility-on-demand systems. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1382–1389, May 2016.

VI. APPENDIX

A. Proof of Theorem 1

Proof: We notice that the uncertainty parameters are involved in both objective functions and constraints. We first use $\max_{F_r \in \mathcal{F}_r} \mathbb{E}(r_i^k)$ substitute r_i^k , $\min_{F_c \in \mathcal{F}_c} \mathbb{E}(c_i^k)$ substitute c_i^k in constraints (7) and (9). It's reasonable due to the idea of minimizing the worst case: any uncertain values of r_i^k and c_i^k should meet the relationship with other decision variables shown in constraints (7) and (9). No matter which probability distribution is selected as the specific distribution to attain the worst case, the simplest worst case for single value of r_i^k and c_i^k is the case that the demand is really large that attains the maximal possible demand value while the supply is very small that attains the minimal possible supply value. And $\mathbb{E}(r_i^k)(\mathbb{E}(c_i^k))$ is the probability-weighted average of all its possible values. Then we transfer all constraints into functional formats as below:

$$\begin{aligned}
 f_{D_i^k} &= \max_{F_r \in \mathcal{F}_r} \mathbb{E}(r_i^k) - l_i^k S_i^k - (D_i^k)^2 = 0, \\
 f_{U_i^k} &= \max_{F_r \in \mathcal{F}_r} \mathbb{E}(r_i^k) - h_i^k S_i^k + (U_i^k)^2 = 0, \\
 f_{S_i^k} &= -S_i^k + X_i^k + V_i^k = 0, \quad k = 1, \dots, \tau, \\
 f_{V_i^{k+1}} &= -V_i^{k+1} + \sum_{j=1}^N P_{vji}^k S_j^k + \sum_{j=1}^N Q_{vji}^k O_j^k \\
 &\quad - \max_{F_c \in \mathcal{F}_c} \mathbb{E}(-c_i^k) = 0, \\
 f_{O_i^{k+1}} &= -O_i^{k+1} + \sum_{j=1}^N P_{oji}^k S_j^k + \sum_{j=1}^N Q_{oji}^k O_j^k = 0, \\
 f_{L_i^{k+1}} &= -L_i^{k+1} + Y_i^k + \sum_{j=1}^N P_{lji}^k S_j^k = 0, \\
 L_i^k &> 0, \quad k = 1, \dots, \tau - 1, \quad S_i^k > 0, \quad k = 1, \dots, \tau.
 \end{aligned} \tag{15}$$

Let $f_D = [f_{D_1^1}, f_{D_1^2}, \dots, f_{D_1^\tau}, \dots, f_{D_N^1}, \dots, f_{D_N^\tau}]^T \in \mathbb{R}^{N\tau}$ be a constraint function vector, for $i = 1, \dots, N, k = 1, \dots, \tau - 1$, and f_U, f_S, f_V, f_O, f_L have the same definition but for computational convenient, if one's dimension is less than $N\tau$, we add 0 in corresponding missing positions to complete its dimension. For the primal maximization problem $\max_{F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c} \mathbb{E}[J_D + \theta J_E]$, s.t (15), its associated Lagrange dual problem (16) can obtain its best upper bound since the strong duality holds. The primal objective function only contains uncertainty

parameter c and is concave over c because it's a linear function of c when given other decision variables. The constraints are also all linear in r and c . When the primal problem is in this case, we usually have strong duality, and Slater's theorem also states that strong duality holds according to [1]. More context about duality are discussed in [1].

$$\begin{aligned} & \min_{\lambda \geq 0} \max_{F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c} J_{dual}, \\ J_{dual} = & \mathbb{E}[J_D + \theta J_E] - (\lambda_U^T f_U + \lambda_S^T f_S + \lambda_V^T f_V \\ & + \lambda_O^T f_O + \lambda_L^T f_L + \lambda_s^T S^{1:\tau} + \lambda_l^T L^{2:\tau}) \end{aligned} \quad (16)$$

$\lambda_U^T, \lambda_S^T, \lambda_V^T, \lambda_O^T, \lambda_L^T, \lambda_s^T, \lambda_l^T$ are corresponding Lagrange multipliers and λ is defined as a vector combined by all these Lagrange multipliers. We have $\frac{c_i^k}{(Y_i^k)^a} \geq 0$ and $c_i^k \geq 0$ by the definitions of J_E in (6) and the Queuing model, then for any vector $Z \in \mathbb{R}^{N\tau}$, $Z = [z_1^1, z_2^1, \dots, z_1^\tau, z_2^\tau, \dots, z_{N\tau}^\tau]^T$ that satisfies $0 < \frac{1}{(Y_i^k)^a} \leq z_i^k$, we also have

$$0 \leq \sum_{k=1}^\tau \sum_{i=1}^N \frac{c_i^k}{(Y_i^k)^a} \leq Z^T c,$$

and the second inequality strictly holds when all $\frac{r_i^k}{(Y_i^k)^a} = z_i^k$, for $i = 1, \dots, N$, $k = 1, \dots, \tau$. The constraints of problem (16) are independent of c , hence, for any c , the minmax problem (16) is equivalent to

$$\begin{aligned} & \min_{\lambda \geq 0} \max_{F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c} J'_{dual} \\ & \text{s.t. } \frac{1}{(Y_i^k)^a} \leq z_i^k, \quad Z \in \mathbb{R}^{N\tau} \end{aligned} \quad (17)$$

where

$$\begin{aligned} J'_{dual} = & \mathbb{E}[J_D + \theta Z^T c] - (\lambda_U^T f_U + \lambda_S^T f_S + \lambda_V^T f_V \\ & + \lambda_O^T f_O + \lambda_L^T f_L + \lambda_s^T S^{1:\tau} + \lambda_l^T L^{2:\tau}) \end{aligned} \quad (18)$$

In the dual problem's objective function J'_{dual} , not all parts contain uncertainty parameters. We can separate J'_{dual} into three parts

$$\begin{aligned} H_r = & -(\lambda_U^T + \lambda_D^T)r, \quad H_c = \theta J_E - \lambda_V^T c, \\ H_o = & J'_{dual} - \mathbb{E}[H_c + H_r], \end{aligned} \quad (19)$$

Where only H_r contains all r , H_c contains all c . H_o can be put as a deterministic value when given other decision variables. So we can turn to consider the following maximization problem

$$\max_{r \sim F_r, c \sim F_c, F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c} \mathbb{E}[H_r + H_c]. \quad (20)$$

Since r and c are independent, problem (20) equals the separated maximization problem

$$\max_{r \sim F_r, F_r \in \mathcal{F}_r} \mathbb{E}[H_r] + \max_{c \sim F_c, F_c \in \mathcal{F}_c} \mathbb{E}[H_c]. \quad (21)$$

Problem (21) satisfies the conditions of Lemma 1 in [2], and the maximum expectation value of $H_r + H_c$ for any possible $r \sim F_r, c \sim F_c$ where $F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c$ equals the optimal value of the problem

$$\begin{aligned} & \min_{\substack{Q_r, q_r, v_r, t_r; \\ Q_c, q_c, v_c, t_c}} v_r + t_r + v_c + t_c \\ & \text{s.t. } v_r \geq H_r - r^T Q_r r - r^T q_r, \\ & t_r \geq (\hat{\gamma}_{2r} \hat{\Sigma}_r + \hat{r} \hat{r}^T) \cdot Q_r + \hat{r}^T q_r \\ & \quad + \sqrt{\hat{\gamma}_{1r}} \|\hat{\Sigma}_r^{1/2} (q_r + 2Q_r \hat{r})\|_2, \quad (22) \\ & v_c \geq H_c - c^T Q_c c - c^T q_c, \\ & t_c \geq (\hat{\gamma}_{2c} \hat{\Sigma}_c + \hat{c} \hat{c}^T) \cdot Q_c + \hat{c}^T q_c \\ & \quad + \sqrt{\hat{\gamma}_{1c}} \|\hat{\Sigma}_c^{1/2} (q_c + 2Q_c \hat{c})\|_2, \\ & Q_r, Q_c \succeq 0. \end{aligned}$$

Note that the first and third constraints about v_r and v_c is equivalent to $v_r \geq f_r(r^*)$ and $v_c \geq f_c(c^*)$ where $f_r(r^*)(f_c(c^*))$ is the optimal value of the following problem

$$\begin{aligned} & \max_r H_r - r^T Q_r r - r^T q_r \quad \text{s.t. } r \geq 0, \\ & \max_c H_c - c^T Q_c c - c^T q_c \quad \text{s.t. } c \geq 0. \end{aligned} \quad (23)$$

Since Q_r and Q_c are positive semi-defined, $H_r(H_c)$ is a linear function over $r(c)$, problem (23) is convex. Solving this problem by taking partial derivative over $r(c)$ without constraints, we have:

$$\begin{aligned} v_r & \geq \frac{1}{4} (q_r + \lambda_U + \lambda_D)^T Q_r^{-1} (q_r + \lambda_U + \lambda_D) \\ v_c & \geq \frac{1}{4} (q_c + \lambda_V - Z)^T Q_c^{-1} (q_c + \lambda_V - Z) \end{aligned} \quad (24)$$

By Schur complement, the above constraints are

$$\begin{aligned} & \begin{bmatrix} v_r & \frac{1}{2} (q_r + \lambda_U + \lambda_D)^T \\ \frac{1}{2} (q_r + \lambda_U + \lambda_D) & Q_r \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} v_c & \frac{1}{2} (q_c + \lambda_V - Z)^T \\ \frac{1}{2} (q_c + \lambda_V - Z) & Q_c \end{bmatrix} \succeq 0, \end{aligned} \quad (25)$$

But this equivalent constraints are under the conditions of no constraints in problem (23). We still have another two constraints for $v_r(v_c)$ that $v_r \geq f_r(r=0) = 0$ and $v_c \geq 0$. Then use the fact that min-min operations can be performed jointly and combine all constraints we can reformulate problem (16) as (14).